

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-259851  
 (43)Date of publication of application : 22.09.2000

(51)Int.Cl. G06T 11/80  
 G06F 3/16  
 G06K 9/00  
 G06K 9/62  
 G10L 15/00  
 G10L 15/28

(21)Application number : 11-062550 (71)Applicant : FUJI XEROX CO LTD  
 (22)Date of filing : 10.03.1999 (72)Inventor : TERAYAMA HIROYUKI

## (54) DEVICE AND METHOD FOR CONVERTING DYNAMIC MODEL

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To convert a model shown in the data of a prescribed form into another model by recognizing the inputted voice data to convert them into a language or an instruction and then converting the attribute or procedure of the image data according to the input time relation between the voice and image data.

**SOLUTION:** The handwritten data or voice data are inputted via a handwritten image input part 3 or a voice input part 5. The inputted handwritten images and voice data are recognized at a handwritten image recognizing part 4 and a voice recognizing part 6 respectively. Then a necessary model is generated from the recognition results of both parts 4 and 6. Two models to be collated with each other are inputted or the features of the time interval, spatial distance, intersection, contact, shape, etc., which were referred to are turned into patterns. These produced patterns are retrieved from those stored in a knowledge base 8. Then the models of the voice or handwritten data and the conversion candidate data are converted according to the rules which are registered on the base 8. Thus, this processing is over.

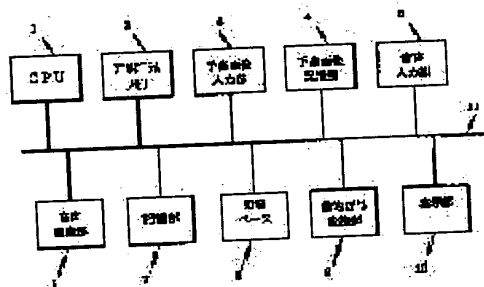


図1 本発明の装置の構成図

## LEGAL STATUS

[Date of request for examination]  
 [Date of sending the examiner's decision of rejection]  
 [Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]  
 [Date of final disposal for application]  
 [Patent number]  
 [Date of registration]

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2000-259851

(P2000-259851A)

(43)公開日 平成12年9月22日(2000.9.22)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	テマコード*(参考)
G 0 6 T 11/80		G 0 6 F 15/62	3 2 0 A 5 B 0 5 0
G 0 6 F 3/16	3 2 0	3/16	3 2 0 A 5 B 0 6 4
G 0 6 K 9/00		G 0 6 K 9/00	S 5 D 0 1 5
9/62		9/62	G
G 1 0 L 15/00		G 1 0 L 3/00	5 5 1 P

審査請求 未請求 請求項の数2 O L (全 7 頁) 最終頁に続く

(21)出願番号 特願平11-62550

(22)出願日 平成11年3月10日(1999.3.10)

(71)出願人 000005496

富士ゼロックス株式会社

東京都港区赤坂二丁目17番22号

(72)発明者 寺山 洋之

神奈川県足柄上郡中井町境430 グリーン

テクなかい 富士ゼロックス株式会社内

(74)代理人 100086531

弁理士 澤田 俊夫

Fターム(参考) 5B050 BA19 CA07 DA09 EA21 FA02

FA10 GA06

5B064 AA07 AB04 BA06

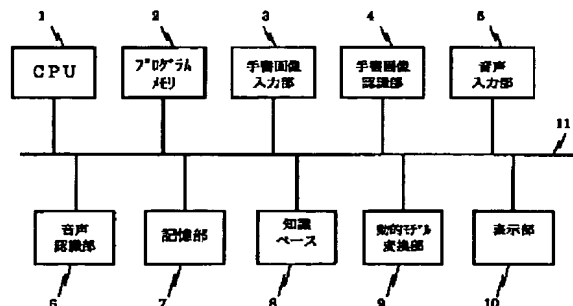
5D015 AA03 BB01

(54)【発明の名称】 動的モデル変換装置及び方法

(57)【要約】

【課題】 既に入力されているようなデータのモデルを、手書きによって加筆される画像によっても、音声によっても、別のモデルに動的に変換することができる動的モデル変換装置及び方法を提供する。

【解決手段】 動的モデル変換装置は、所定の画像空間内に第1の画像データを入力する手段と、音声データを入力する手段と、入力された第1の音声データを認識して言語又は命令に変換する音声認識手段と、第1の音声データと第1の画像データが入力された時間的關係に基づいて第1の画像データが持つ属性又は手続きを変換するデータ・モデル変換手段とを備える。一連の画像データを認識し処理する際に、既に入力された様々なデータのモデルを、手書きによって加筆される画像によっても、音声によっても、適切にモデルを認識し、変換することができる。また、知識ベースを交換することによって、操作者に応じた画像認識と動的モデル変換を容易に実現することができる。



↑  
動的モデル変換装置100

1

## 【特許請求の範囲】

【請求項1】 所定形式のデータで表されるモデルを他のモデルに変換する動的モデル変換装置であって、所定の画像空間内に第1の画像データを入力する手段と、

音声データを入力する手段と、

入力された第1の音声データを認識して言語又は命令に変換する音声認識手段と、

第1の音声データと第1の画像データが入力された時間的関係に基づいて第1の画像データが持つ属性又は手続きを変換するデータ・モデル変換手段と、を具備すること

を特徴とする動的モデル変換装置。

【請求項2】 所定形式のデータで表されるモデルを他のモデルに変換する動的モデル変換方法であって、(a) 所定の画像空間内に第1の画像データを入力するステップと、(b) 第1の音声データを入力するステップと、

(c) 第1の音声データを認識して言語又は命令に変換するステップと、(d) 第1の音声データと第1の画像データが入力された時間的関係に基づいて第1の画像データが持つ属性又は手続きを変換するステップと、を具備することを特徴とする動的モデル変換方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、所定形式のデータで表されるモデルを他のモデルに変換する動的モデル変換装置及び方法に係り、特に、手書きデータや音声データ、静止画像データ、動画データ、手書きや音声の認識結果である文字・図形・ジェスチャ（命令）で表されるモデルを他のモデルに動的に変換する動的モデル変換装置及び方法に関する。

【0002】 更に詳しくは、本発明は、既に入力されているようなデータのモデルを、手書きによって加筆される画像によっても、音声によっても、別のモデルに動的に変換する動的モデル変換装置及び方法に関する。

## 【0003】

【従来の技術】 昨今の情報処理技術の発展に伴い、高性能で且つ強力な演算能力を持つ汎用コンピュータ・システムが、各種研究機関や企業内のオフィス、一般家庭へと広汎に普及してきている。

【0004】 この種のコンピュータは、通常、キーボードやマウス、ディスプレイなどのように、ユーザからの入力やユーザへの情報提示を実現するための「ユーザ・インターフェース」を備えている。ユーザは、キーボードを介してキャラクタ・ベースのコマンド入力を行なうことができる。また、ビットマップ形式の表示機能によって実現されるGUI (Graphical User Interface) 環境下のデスクトップ画面上では、ユーザは必ずしもコマンドをキー入力する必要はなく、関心のあるオブジェクトに対して直感的にマウス操作（クリック、ドラッグ、ドロップなど）することでシ

2

ステムに所望の動作を実行させることが許されている。

【0005】 マウスによる入力操作はキーボード入力よりも直感的で分かり易く、ユーザ・フレンドリである。しかしながら、マウスにしてもキーボードにしても、専門家ではない一般的な人間の感覚に対して依然として不自然さが残り、入力操作が定着するにはある程度以上の時間と労力を要する。このため、より自然な感覚でユーザ入力を実現するユーザ・インターフェースが期待され、これまでも数多くの技術が提案され開発されている。1つの例は音声入力であり、他の例は手書き入力である。

【0006】 このうち手書き入力は、タブレットなどの略平坦な入力面上でペンやスタイラスなどで手書き入力された文字や図形をシステムが自動的に認識して、システムへのコマンドやデータの入力として扱うことによって実現される。

【0007】 従来、タブレットのような入力装置から手書き入力されたものを認識させる装置では、手書き入力されたデータが文字であればパターン・マッチングなどにより文字として認識され対応する文字コードに変換される。また手書きデータが図形や線画であれば、夫々図形データ又は線画データとして、ベクトル・データ又はラスター・データに変換して扱われる。また、手書きデータがジェスチャとして認識された場合には、対応する命令などに変換するして用いられる。言い換えれば、手書きデータという形式で入力されたモデルは、他のモデルに変換されてシステム内で扱われるというのが一般的であった。

【0008】 例えば、特開平6-111073号公報に開示された手書き情報認識装置は、表示一体型タブレットに入力された筆跡座標、あるいは文書情報記憶装置に記憶されている既入力文書情報に対して、文字らしさ、図形らしさ、編集用記号らしさなど、夫々のカテゴリを識別するための評価項目に従って各カテゴリ毎の確信度を付し、最も高得点か又は閾値以上の得点を持つカテゴリを入力情報のカテゴリとして識別するようになっている。

【0009】 また、画像で表されるモデルを変換する幾つかの技術についても、既に幾つかの提案がなされている。例えば、特開平7-110872号公報では、複数の2次元画像データから3次元データのモデルに変換するソリッド・モデル合成装置及び合成方法について開示している。すなわち、同合成装置は、入力した2次元図形データを平面図データと断面図データとに分類し、平面図データから閉領域と各閉領域間の包含関係を検出するとともに、検出された閉領域の属性を検出して、同じ属性を持つ閉領域をグループ化する。さらに、検出された閉領域のデータに基づいて断面図データを分割して、分割された断面図データに基づいて閉領域を立体化するとともに、立体化された閉領域と同じグループに属する

その他の閉領域についても同じ方法で立体化する。そして、各閉領域間の包含関係に基づいて集合演算により各閉領域を組み合わせ、ソリッド・モデルを合成する。

【0010】また本出願人に既に譲渡されている特願平10-266861号明細書（社内整理番号：FE98-00181）は、手書きデータを含む画像データのモデルを他のモデルに動的に変換する動的モデル変換装置について開示している。すなわち、同動的モデル変換装置は、入力された手書きデータの近傍にある近傍データのリストを作成し、このリストから近傍データを取り出して、手書きデータとの関係をパターン化して、このパターンと知識ベースに格納されているパターンとの照合を行ない、動的にモデルを変換するようになっている。

【0011】しかしながら、上記の各特許出願において開示された技術はいずれも、画像データだけに基づく変換であるため、入力が煩わしいという問題がある。

【0012】

【発明が解決しようとする課題】本発明の目的は、所定形式のデータで表されるモデルを他のモデルに変換することができる、優れた動的モデル変換装置及び方法を提供することにある。

【0013】本発明の更なる目的は、手書きデータや音声データ、静止画像データ、動画画像データ、手書きや音声の認識結果である文字・図形・ジェスチャ（命令）で表されるモデルを他のモデルに動的に変換することができる、優れた動的モデル変換装置及び方法を提供することにある。

【0014】本発明の更なる目的は、既に入力されているようなデータのモデルを、手書きによって加筆される画像によっても、音声によっても、別のモデルに動的に変換することができる、優れた動的モデル変換装置及び方法を提供することにある。

【0015】

【課題を解決するための手段及び作用】本発明は、上記課題を参酌してなされたものであり、その第1の側面は、所定形式のデータで表されるモデルを他のモデルに変換する動的モデル変換装置であって、所定の画像空間内に第1の画像データを入力する手段と、音声データを入力する手段と、入力された第1の音声データを認識して言語又は命令に変換する音声認識手段と、第1の音声データと第1の画像データが入力された時間的關係に基づいて第1の画像データが持つ属性又は手続きを変換するデータ・モデル変換手段とを具備することを特徴とする動的モデル変換装置である。

【0016】また、本発明の第2の側面は、所定形式のデータで表されるモデルを他のモデルに変換する動的モデル変換方法であって、（a）所定の画像空間内に第1の画像データを入力するステップと、（b）第1の音声データを入力するステップと、（c）第1の音声データを認識して言語又は命令に変換するステップと、（d）

第1の音声データと第1の画像データが入力された時間的關係に基づいて第1の画像データが持つ属性又は手続きを変換するステップと、を具備することを特徴とする動的モデル変換方法である。

【0017】本発明に係る動的モデル変換装置によれば、一連の画像データを認識し処理する際に、既に入力された様々なデータのモデルを、手書きによって加筆される画像によっても、音声によっても、適切にモデルを認識し、変換することができる。

【0018】また、知識ベースを交換することによって、操作者に応じた画像認識と動的モデル変換を容易に実現することができる。

【0019】本発明のさらに他の目的、特徴や利点は、後述する本発明の実施例や添付する図面に基づくより詳細な説明によって明らかになるであろう。

【0020】

【発明の実施の形態】以下、図面を参照しながら本発明の実施例を詳解する。

【0021】図1には、本発明の実施に供される動的モデル変換装置100のハードウェア構成を模式的に示している。本実施例に係る動的モデル変換装置100は、画像処理という業務に適用したものである。以下、各部について説明する。

【0022】CPU（Central Processing Unit）1は、動的モデル変換装置100全体の動作を統括的に制御するためのメイン・コントローラである。一般に、CPU1は、オペレーティング・システム（OS）の制御下で各種プログラム・コードを実行し、本実施例においては動的モデル変換に関わる各種処理のための演算や論理判断などを行なう。

【0023】CPU1は、「バス」と呼ばれる共通信号伝送路11を介して各種周辺装置と相互接続されている。

【0024】プログラム・メモリ2は、CPU1が実行するプログラム・コードをロードしたり、処理中の作業データを一次書き込むための書き込み可能メモリである。プログラム・メモリ2は、通常、複数のDRAM（Dynamic Random Access Memory）チップで構成される。

【0025】手書き画像入力部3は、例えば「タブレット」のような、画像データのペン入力を受容する装置である。ペン入力指示座標の読み取り方式は、感圧式、静電容量式など、特に限定されない。

【0026】手書き画像認識部4は、手書き画像入力部3から手書き入力された画像データを文字、図形、ジェスチャ等に分類して認識する機能モジュールである。

【0027】音声入力部5は、例えばマイクのような、音声データを入力する装置である。また、音声認識部6は、音声データを認識して、対応する言語や命令に変換する機能モジュールである。

5

【0028】記憶部7は、手書き画像データや音声データ、マルチメディア・データ、CPU1において処理されたデータのモデル、及び、各モデル間の関係を記述したデータを記憶する機能モジュールである。記憶部7の実体は、例えば、ハード・ディスク装置のような大容量の記憶装置、又は、ハード・ディスク内に格納されたファイルである。

【0029】知識ベース8は、入力データのモデルに対するモデル変換を行うための各種の知識を保有するモジュールであり、その実体はハード・ディスク装置のような大容量記憶装置、又はハード・ディスク内に格納されたファイルである。

【0030】ここで言う「知識」とは、具体的には、2以上のモデル間で生じうる関係や、関係が生じた場合のモデル変換手順などを記述した情報である。また、ここで言う「関係」とは、例えば、モデル間の距離や入力的时间間隔、包含、交差、交点や接点の数などのことを指す。これらの知識は、一般的なIF-THEN形式のルールで表現することができる。

【0031】動的モデル変換部9は、手書きデータや音声データと、時間的あるいは空間的に近いデータとの関係から、夫々のモデルを動的に変換する処理を行う機能モジュールである。但し、モデルの動的変換に関する処理手順の詳細については後述する。

【0032】出力部10は、手書き画像データや音声データ、マルチメディアデータ、記憶部7に記憶されたデータなどを画像で表示し、音声データを再生する装置である。具体的には、CRT (Cathode Ray Tube) やLCD (Liquid Crystal Display) のような表示装置と、スピーカのような音声出力装置の組み合わせで構成される。

【0033】なお、上述した動的モデル変換装置100は、動的モデル変換のために特化してデザインされた専用のハードウェア装置として実装される他、ワークステーションやパーソナル・コンピュータと呼ばれる汎用コンピュータ・システム上で所定のアプリケーション・プログラムを実行するという形態でも実現される。アプリケーションの一例は、動的モデル変換用プログラムである。

【0034】次いで、動的モデル変換装置100によるモデルの動的変換処理について説明する。図2には、手書き画像の処理に認識結果を動的に反映させるための動的モデル変換処理の手順をフローチャートの形式で示している。以下、このフローチャートの各ステップについて説明する。

【0035】まず、ステップS101では、手書きデータ又は音声データを、手書き画像入力部3又は音声入力部5を介して入力する。入力された手書き画像は手書き画像認識部4で、音声データは音声認識部6で、夫々認識され、各々の認識結果から必要となるモデルが生成さ

6

れる。なお、手書き画像認識及び音声認識の処理手順自体は周知技術を適用可能なので、本明細書ではこれ以上説明しない。

【0036】次いで、ステップS102では、ステップS101で入力されたデータを基に、時間的或いは空間的に近い1以上のデータとそれらの組み合わせからなる変換候補データ・リストを生成する。

【0037】次いで、ステップS103では、初期のパターン照合対象をセットする。

【0038】次いで、ステップS104では、照合する2つのモデルが入力され又は参照された時間的間隔や空間的な距離、交点、接点、形状などの特徴をパターン化する。

【0039】次いで、ステップS105では、知識ベース8に格納されているパターンの中から、ステップS104で作成されたパターンを検索する。

【0040】ステップS106では、マッチするパターンが知識ベース8中で発見できたか否かをチェックする。判断ブロックS106の結果、マッチするパターンを発見できたならばステップS107へ進み、発見できなかったらステップS108へ進む。

【0041】ステップS107では、知識ベース8に登録されているルールに従って、音声又は手書きデータのモデルや変換候補データのモデル、あるいは音声又は手書きデータと変換候補データの両方のモデルを変換して、この処理フロー全体を終了する。但し、ここで言う「モデルの変換」とは、属性や手続きの追加や変更、削除、モデル自体の複製や削除などを意味する。

【0042】他方、ステップS108では、上記のステップS102において作成された変換候補データ・リストをチェックする。リスト中に未照合の変換候補データが存在する場合は次ステップS109に進み、存在しない場合にはこの処理フロー全体を終了する。

【0043】ステップS109では、変換候補データ・リストから次の変換候補データを取り出し、ステップS104に復帰して上述と同様の処理を繰り返す行なう。

【0044】図3には、モデルを動的に変換する様子を図解している。以下、この図に従って説明する。

【0045】図3④は、手書きされた4角形aを示している。この4角形aは、始点の座標(x1, y1)、幅(w1)、高さ(h1)という各属性値を有するフレーム・モデルである。

【0046】4角形aを手書きする前後に「マトリクス」という文字列を音声入力すると、音声認識部6による認識結果に基づいて、言語データ「マトリクス」を属性として持つモデルが生成される。

【0047】ここで、“4角形aを手書きした前後の適当な時間間隔内に「マトリクス」という音声が入力され認識された場合は4角形aがマトリクスbである”、という旨の知識(ルール)が知識ベース8に予め登録され

7

ているとする。この場合、手書き入力された図形は、音声入力に従って、表に変換される。(図3②)

【0048】図3③は、マトリクスbに対して重なるように線分cが手書きされた状態を示している。線分cは、単独では、始点座標と終点座標という座標情報のみで表されるモデルになる。

【0049】ここで、“マトリクスbの同じ列に重なる線分cは列の分割あるいは列の挿入を表す”、という旨の知識が知識ベース8に予め登録されているものとする。この場合、線分cは音声入力に従ってマトリクスの列に変換される。すなわち、マトリクスの列を容易に増やすことができる訳である。

【0050】同様に、図3④は、知識ベース8に登録された知識に従って、手書きされた線分dによってマトリクスの行が増やされた状態を示している。

【0051】線分dは、単独では、始点座標と終点座標という座標情報のみで表されるモデルになる。”マトリクスbの同じ行に重なる線分dは列の分割あるいは行の挿入を表す”、という旨の知識が知識ベース8に予め登録しておくことにより、線分dは音声入力に従ってマトリクスの行に変換される。すなわち、マトリクスの列も容易に増やすことができる訳である。

【0052】なお、マトリクスbに対する加筆は、上述したようなマトリクス内の行や列の追加以外に、マトリクスの外側に接する線分などで手書きした場合でも、必要な知識が登録されていれば変換が可能であることは言うまでもない。

【0053】〔追補〕以上、特定の実施例を参照しながら、本発明について詳解してきた。しかしながら、本発明の要旨を逸脱しない範囲で当業者が該実施例の修正や代用を成し得ることは自明である。すなわち、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本発明の要旨を判断するためには、冒頭に記載した特許請求の範囲の欄を参酌すべきである。

【0054】

【発明の効果】以上詳記したように、本発明によれば、所定形式のデータで表されるモデルを他のモデルに変換することができる、優れた動的モデル変換装置及び方法を提供することができる。

【0055】また、本発明によれば、手書きデータや音

8

声データ、静止画像データ、動画像データ、手書きや音声の認識結果である文字・図形・ジェスチャ（命令）で表されるモデルを他のモデルに動的に変換することができる、優れた動的モデル変換装置及び方法を提供することができる。

【0056】本発明に係る動的モデル変換装置によれば、一連の画像データを認識し処理する際に、既に入力された様々なデータのモデルを、手書きによって加筆される画像によっても、音声によっても、適切にモデルを認識し、変換することができる。また、知識ベースを交換することによって、操作者に応じた画像認識と動的モデル変換を容易に実現することができる。

【図面の簡単な説明】

【図1】 本発明の実施に供される動的モデル変換装置10のハードウェア構成を模式的に示した図である。

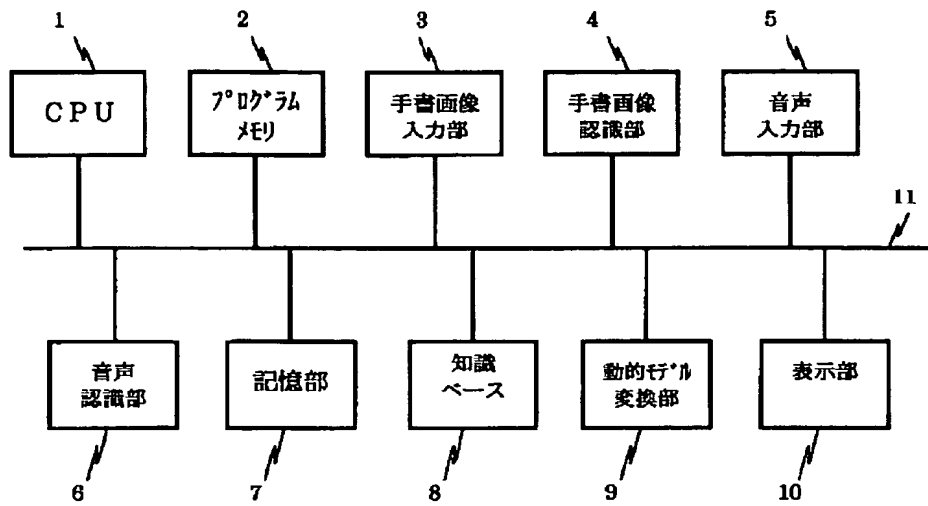
【図2】 動的モデル変換の処理手順を示したフローチャートである。

【図3】 手書き入力された画像のモデルを動的に変換する具体例を図解したものであり、より具体的には、図3①は手書きされた4角形aを表した図、図3②は知識ベース8に登録された知識に従って手書きされた4角形aがマトリクスbに変換された様子を示した図、図3③はマトリクスbの列に対して重なるように手書き入力された線分cが知識ベース8に登録された知識に従ってマトリクスの列に変換された状態を示した図、図3④はマトリクスbの行に対して重なるように手書き入力された線分dが知識ベース8に登録された知識に従ってマトリクスの行に変換された状態を示した図である。

【符号の説明】

- 1…CPU
- 2…プログラム・メモリ
- 3…手書き画像入力部
- 4…手書き画像認識部
- 5…音声入力部
- 6…音声認識部
- 7…記憶部
- 8…知識ベース
- 9…動的モデル変換部
- 10…表示部
- 11…バス
- 100…動的モデル変換装置

【図1】



↑  
動的モデル変換装置 100

【図3】

① 既入力 of 四角形



四角形はフレームモデル

② 状態①で「マトリクス」と音声入力



デフォルトでは2×2のマトリクスに変換

③ 状態②に横線を加筆



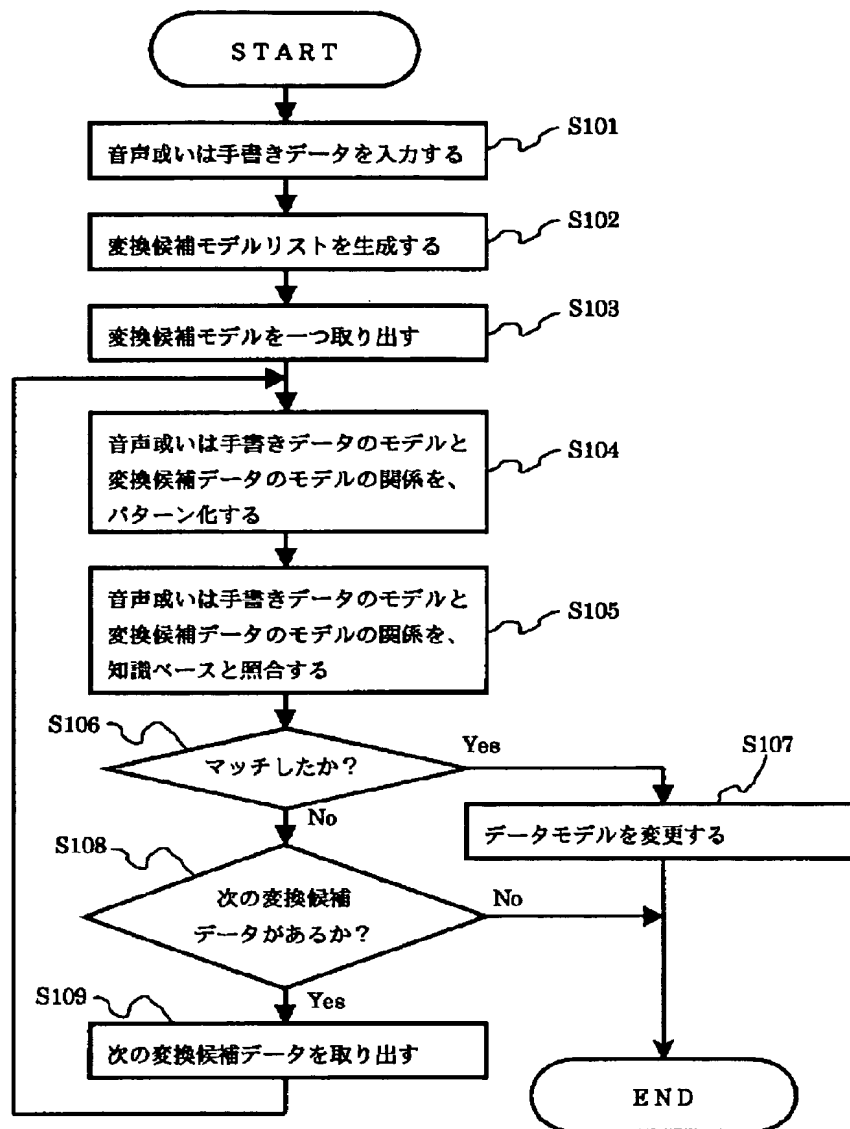
2×3のマトリクスに変換

④ 状態③に横線を加筆



3×3のマトリクスに変換

【図2】



フロントページの続き

(51) Int. Cl. 7

識別記号

F I

テーマコード\* (参考)

G 1 0 L 15/28